Preserving Meaning of Evidence from Evolving Systems

Hannes Spichiger (HSLU), hannes.spichiger@hslu.ch
Frank Adelstein (Hexordia), frank_adelstein@hexordia.com

HSLU Hochschule Luzern

HEXORDIA

Origin of This Work: DFRWS-EU 2024 Post-Rodeo

- Rodeo ended at midnight, then an hour-long conversation afterwards.
- 6 months of further discussion comparing aspects of traditional forensics to digital forensics.
- Writing and rewriting
- DFRWS is a place for collaboration!

The Three Bandito Rodeo Wranglers (Razvan, Javier, and Ricardo)

The Main Point

- This is NOT about:

  “Everything you’re doing is wrong. Your conclusions are invalid.”

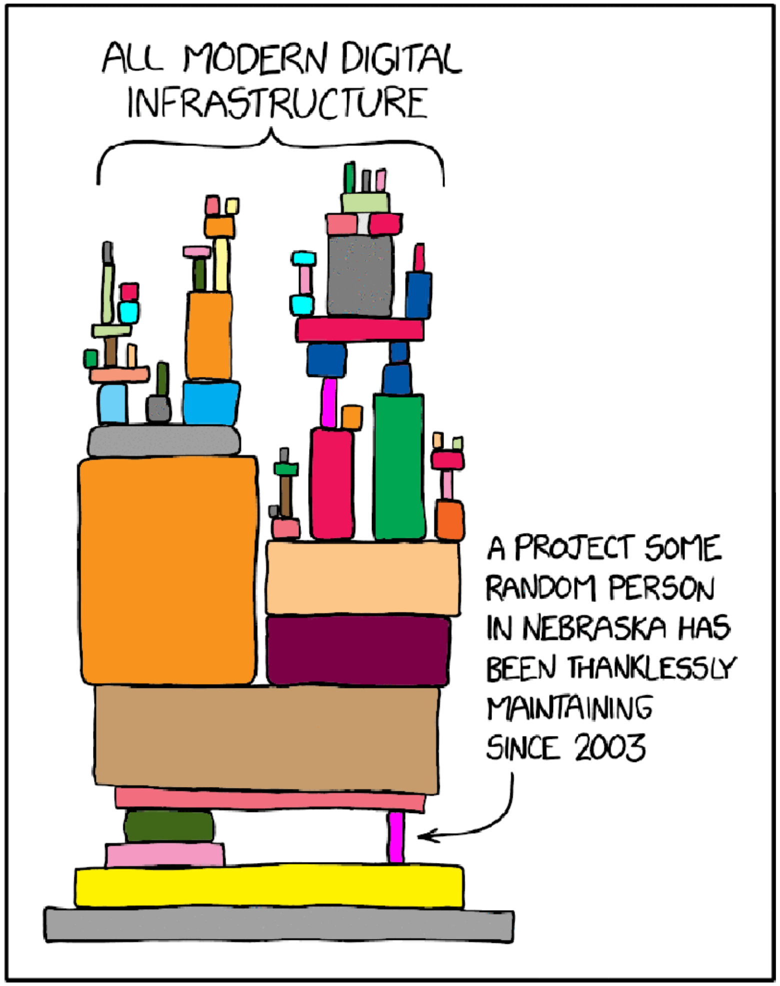
- Conclusions reached through proper DF methods are and remain valid.
- This IS about:

  The pace of change will make it more challenging to reach some conclusions

From Distributed Systems to Evolving Systems

Modern client/server distributed systems change much faster than they did 10 or 20 years ago.

This was the model of software infrastructure just 5 years ago…


ALL MODERN DIGITAL INFRASTRUCTURE
A PROJECT SOME RANDOM PERSON IN NEBRASKA HAS BEEN THANKLESSLY MAINTAINING SINCE 2003

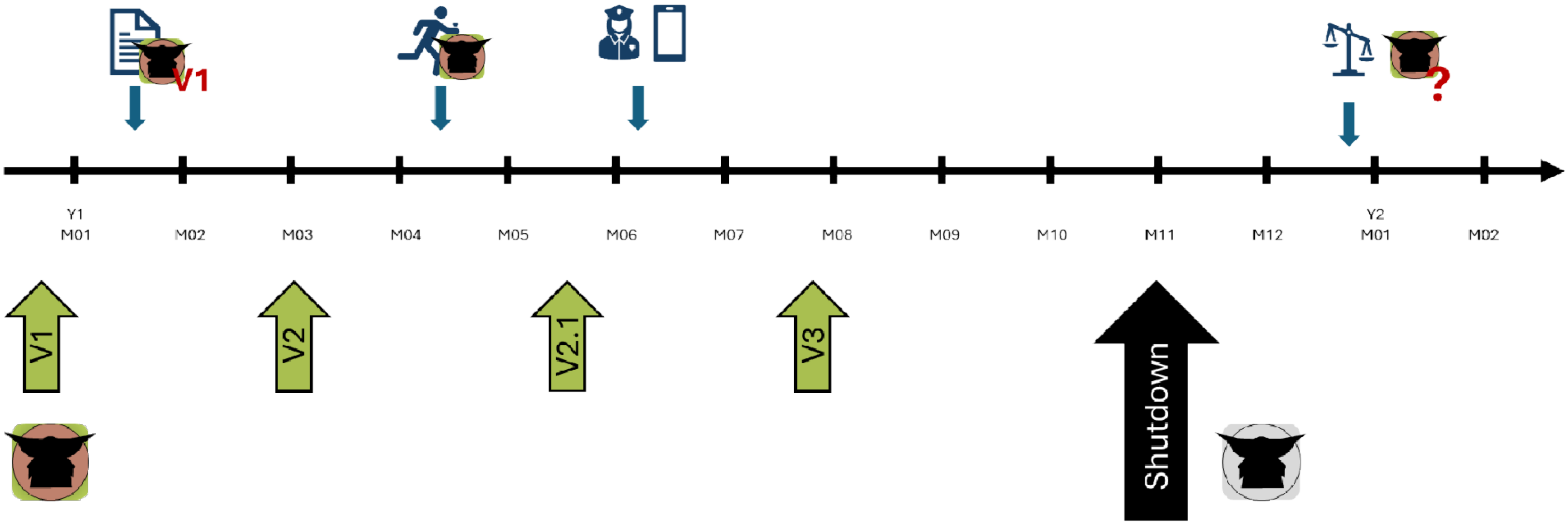
https://xkcd.com/2347

From Distributed Systems to Evolving Systems

Evolving Systems: highly dynamic distributed systems where components of both the server and client-side change often, which can impact the stability of reference data.

Reference Data: data used as a standard to classify or interpret trace data acquired from digital devices.

Investigating a mobile application



Impact on Investigations

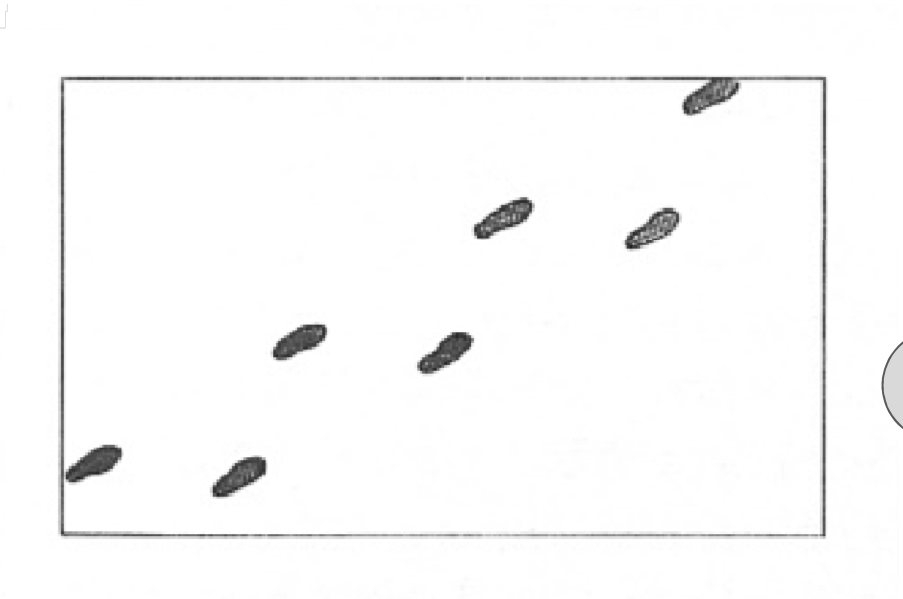
Does this mean you can’t say anything about the data on the phone?

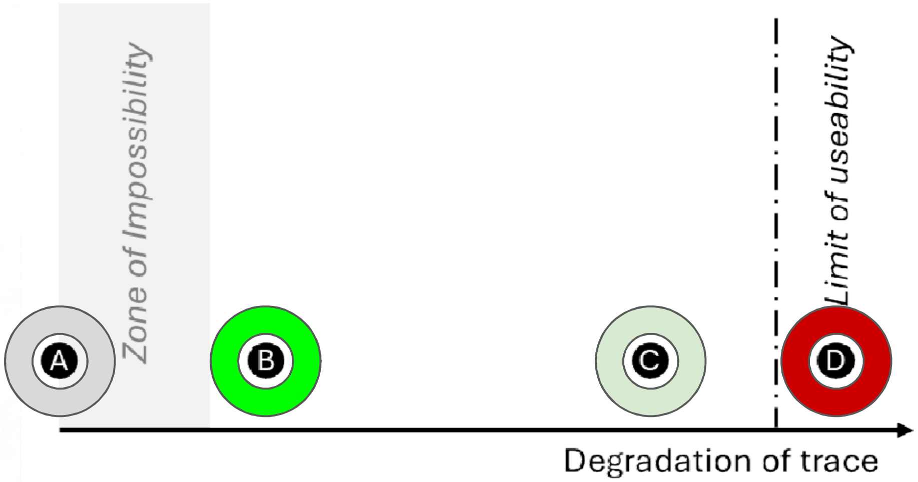
Not at all!

But it does mean that the level of uncertainty has increased.

The *meaning* of the evidence has degraded despite that it is perfectly preserved.

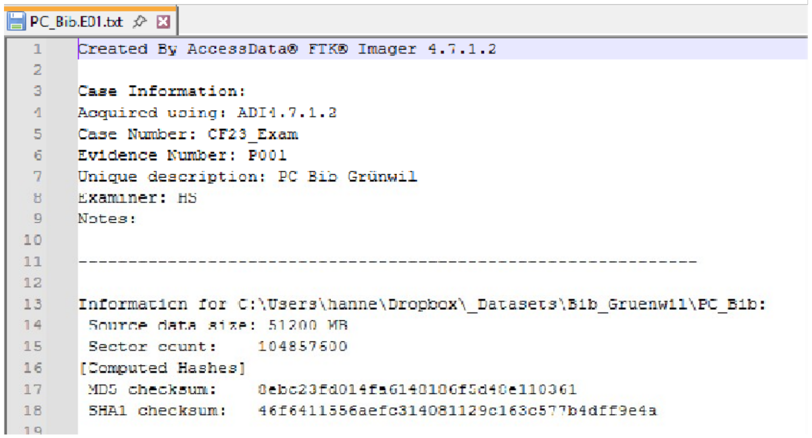
Trace Degradation (Traditional)





Kind, S.S., 1994. Crime investigation and the criminal trial: a three chapter paradigm of evidence. Journal of the Forensic Science Society 34, 155–164. *https://doi.org/10.1016/S0015-7368(94)72908-X*

Trace Degradation (Digital)



```
Created By AccessData® FTK® Imager 4.7.1.2

Case Information:
Acquired using: ADI4.7.1.2
Case Number: CF23_Exam
Evidence Number: P001
Unique description: PC Bib Grünwil
Examiner: HS
Notes:

----------------------------------------------------------------

Information for C:\Users\hanne\Dropbox\_Datasets\Bib_Gruenwil\PC_Bib:
 Source data size: 51200 MB
 Sector ccunt:     104857600
[Computed Hashes]
 MD5 checksum:     8ebc23fd014fa6140186f5d48e110361
 SHA1 checksum:    46f6411556aefc314081129c163c577b4dff9e4a
```

Image source: https://www.itp.net/acn/cybersecurity/585942-faraday-bags-help-secure-seized-mobile-devices

What About Reference Data

Can reference data decay?

If so, how?

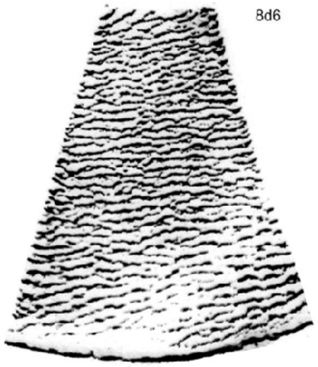
Traditional Forensic Science

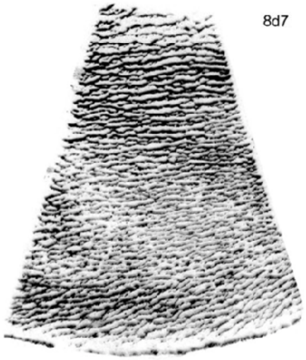
Trace	**Reference**
Fingermarks	Fingerprints
Shoe Marks	Boot soles
Tire tracks	Tires
Barrel marks on bullets	Gun barrel

Reference Degradation (Physical)

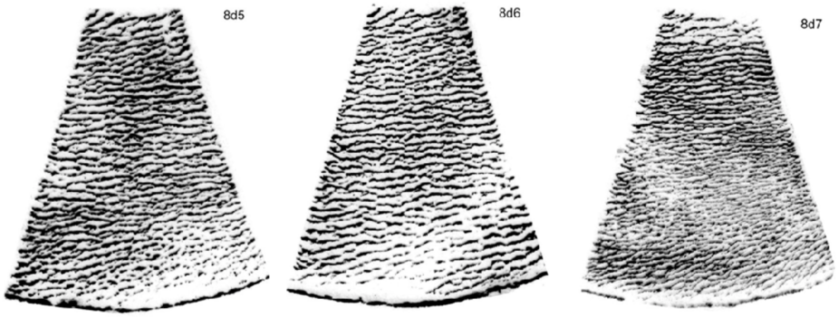
Shoes and tires get worn over time and no longer match the traces

Gun barrels are altered over time by use and no longer create the same patterns

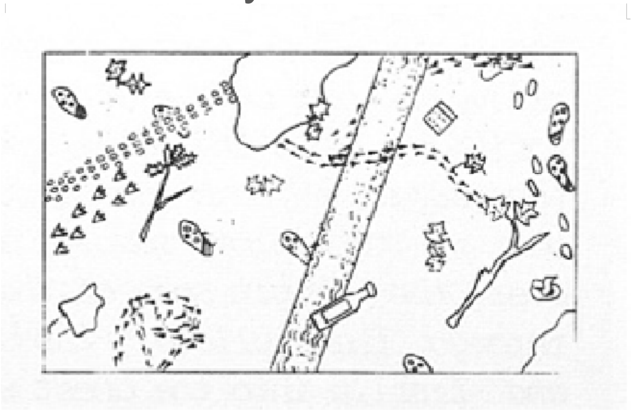




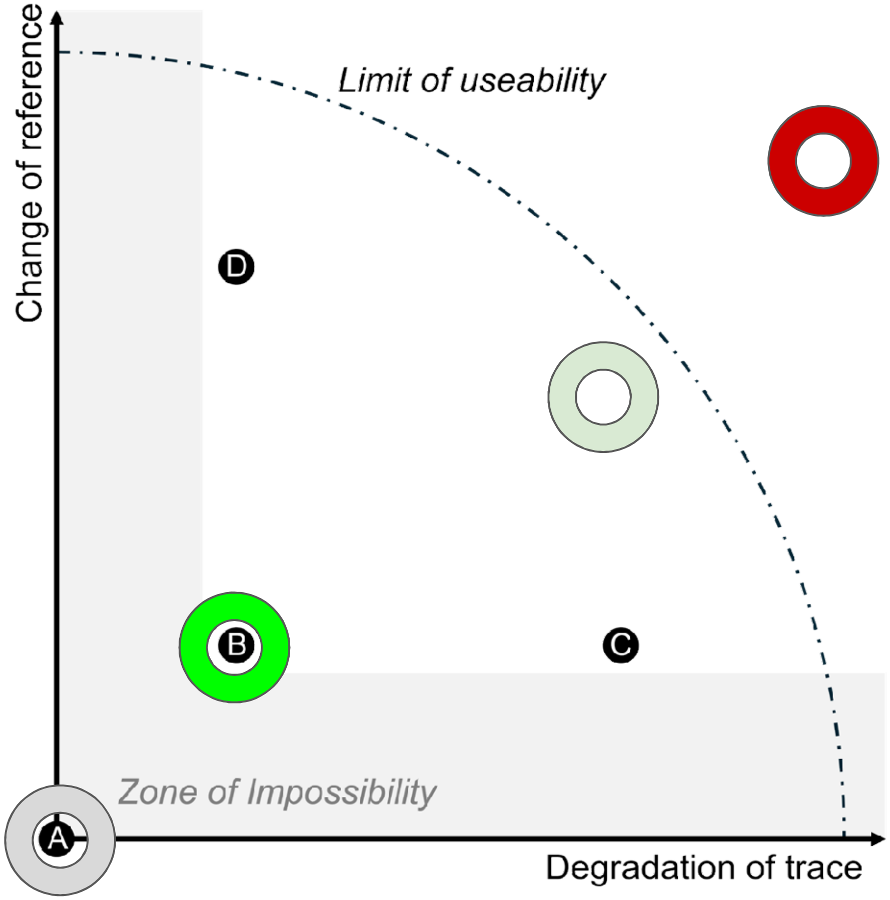
Stauffer, G., 2000. Modèle de Schallamach -Compréhension et description de ce phénomène d'usure et exploitation de cette information dans le cadre de l'examen des traces de semelles, Séminaire de 4e année. Institut de Police Scientifique et de Criminologie. Université de Lausanne, Lausanne.



Wear on Army Boot Sole



Decay of Footprints in the Snow



DF Reference Data

Reference Data: data used as a standard to classify or interpret trace data acquired from digital devices.

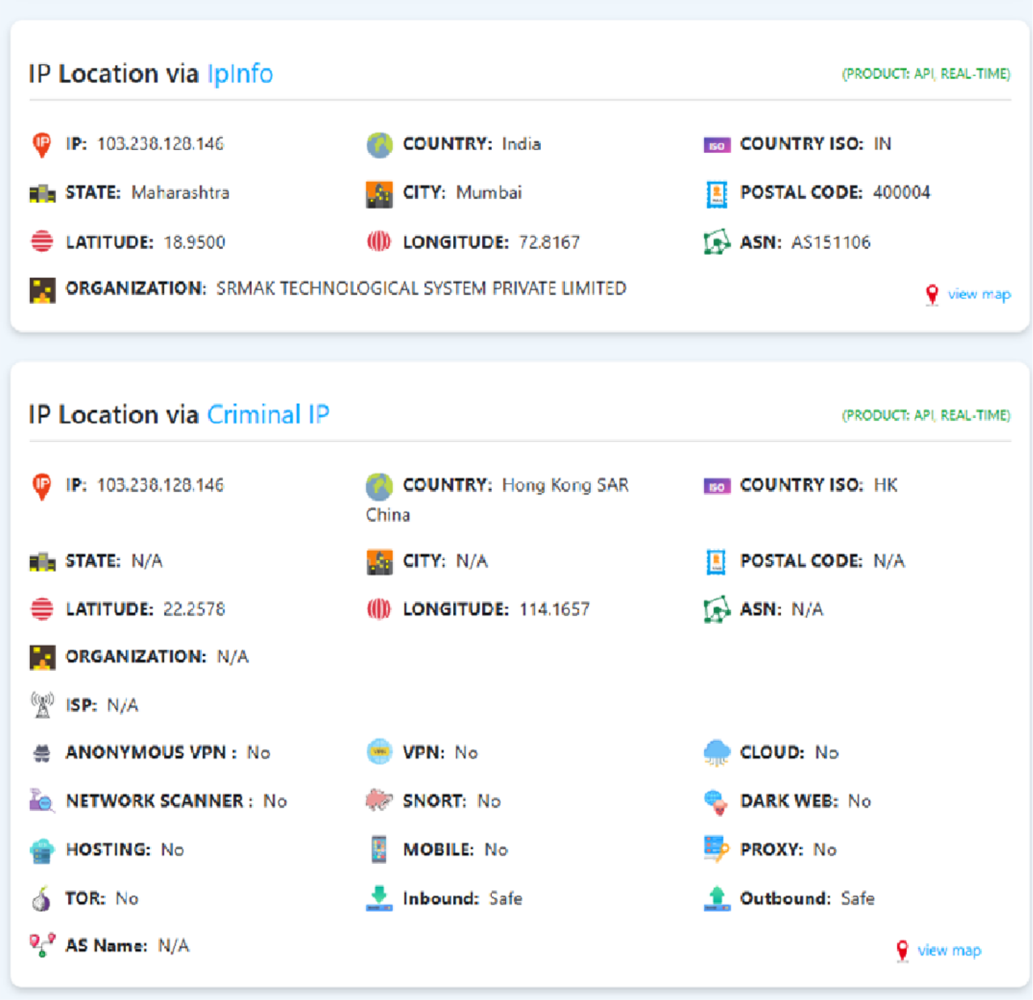
What is the meaning of the following value? 1743683730

Without reference data, it can be very difficult to know how to interpret the raw data value.

A timestamp? => Thursday, April 3, 2025 12:35:30 PM GMT (our talk time)

An IP-Address? => 103.238.128.146

IP-Addresses



IP Location via IpInfo

(PRODUCT: API, REAL-TIME)

IP: 103.238.128.146

COUNTRY: India

COUNTRY ISO: IN

STATE: Maharashtra

CITY: Mumbai

POSTAL CODE: 400004

LATITUDE: 18.9500

LONGITUDE: 72.8167

ASN: AS151106

ORGANIZATION: SRMAK TECHNOLOGICAL SYSTEM PRIVATE LIMITED

view map

IP Location via Criminal IP

(PRODUCT: API, REAL-TIME)

IP: 103.238.128.146

COUNTRY: Hong Kong SAR China

COUNTRY ISO: HK

STATE: N/A

CITY: N/A

POSTAL CODE: N/A

LATITUDE: 22.2578

LONGITUDE: 114.1657

ASN: N/A

ORGANIZATION: N/A

ISP: N/A

ANONYMOUS VPN : No

VPN: No

CLOUD: No

NETWORK SCANNER : No

SNORT: No

DARK WEB: No

HOSTING: No

MOBILE: No

PROXY: No

TOR: No

Inbound: Safe

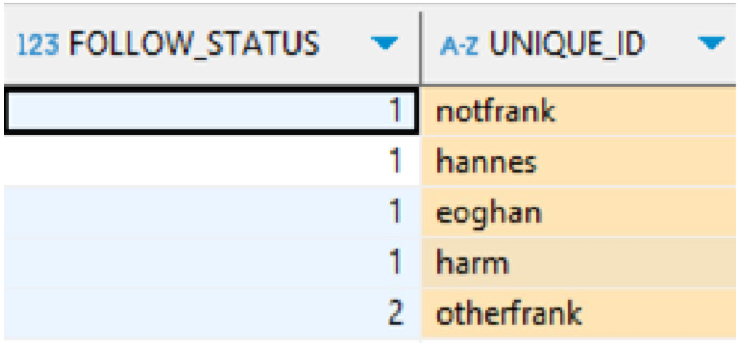
Outbound: Safe

AS Name: N/A

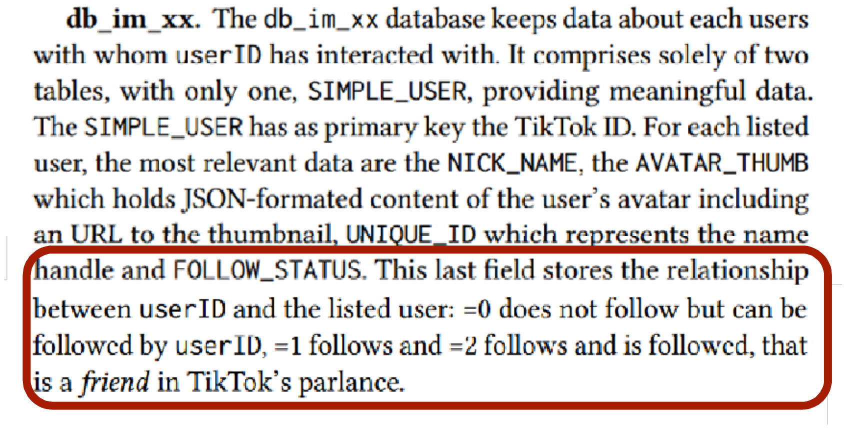
view map

https://iplocation.io

Trace Data is understood through Reference Data



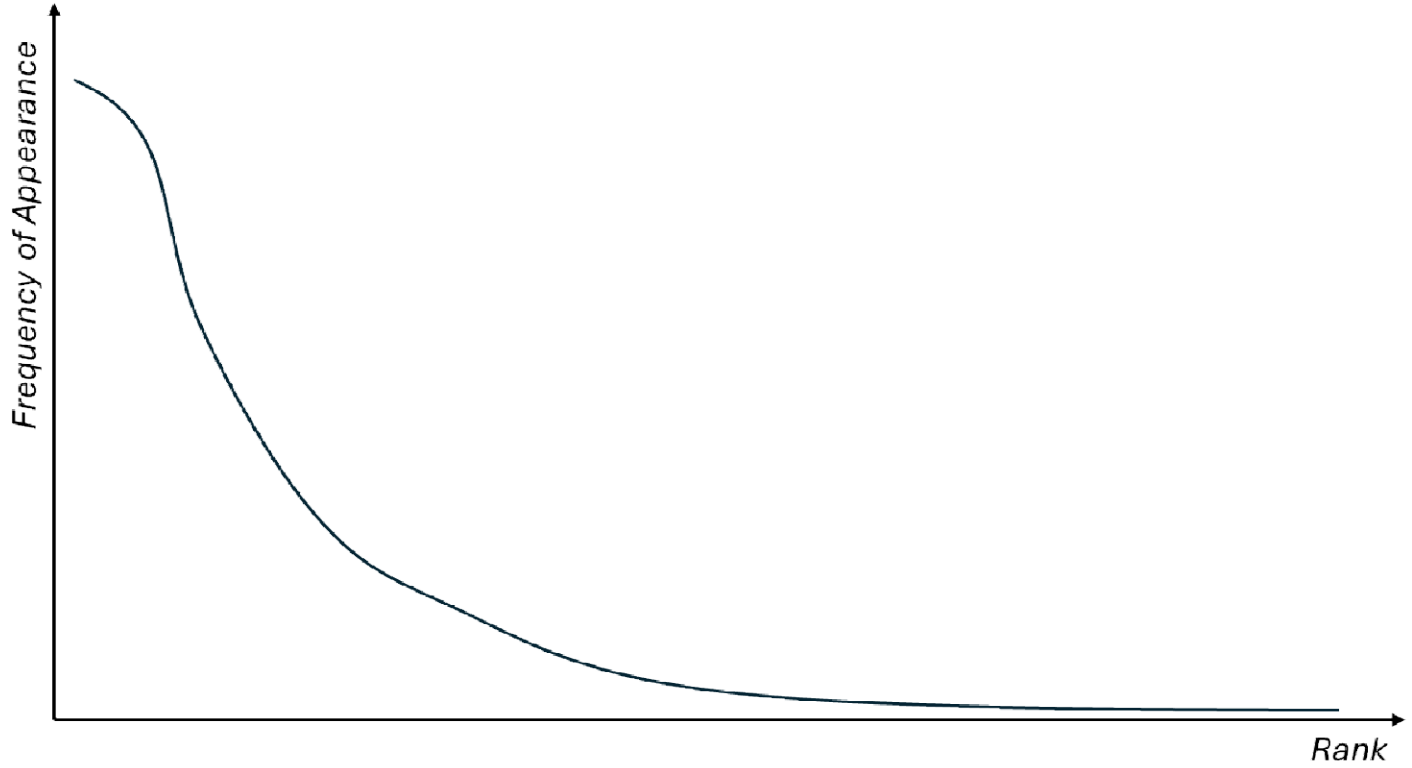
123 FOLLOW_STATUS	A-Z UNIQUE_ID
1	notfrank
1	hannes
1	eoghan
1	harm
2	otherfrank



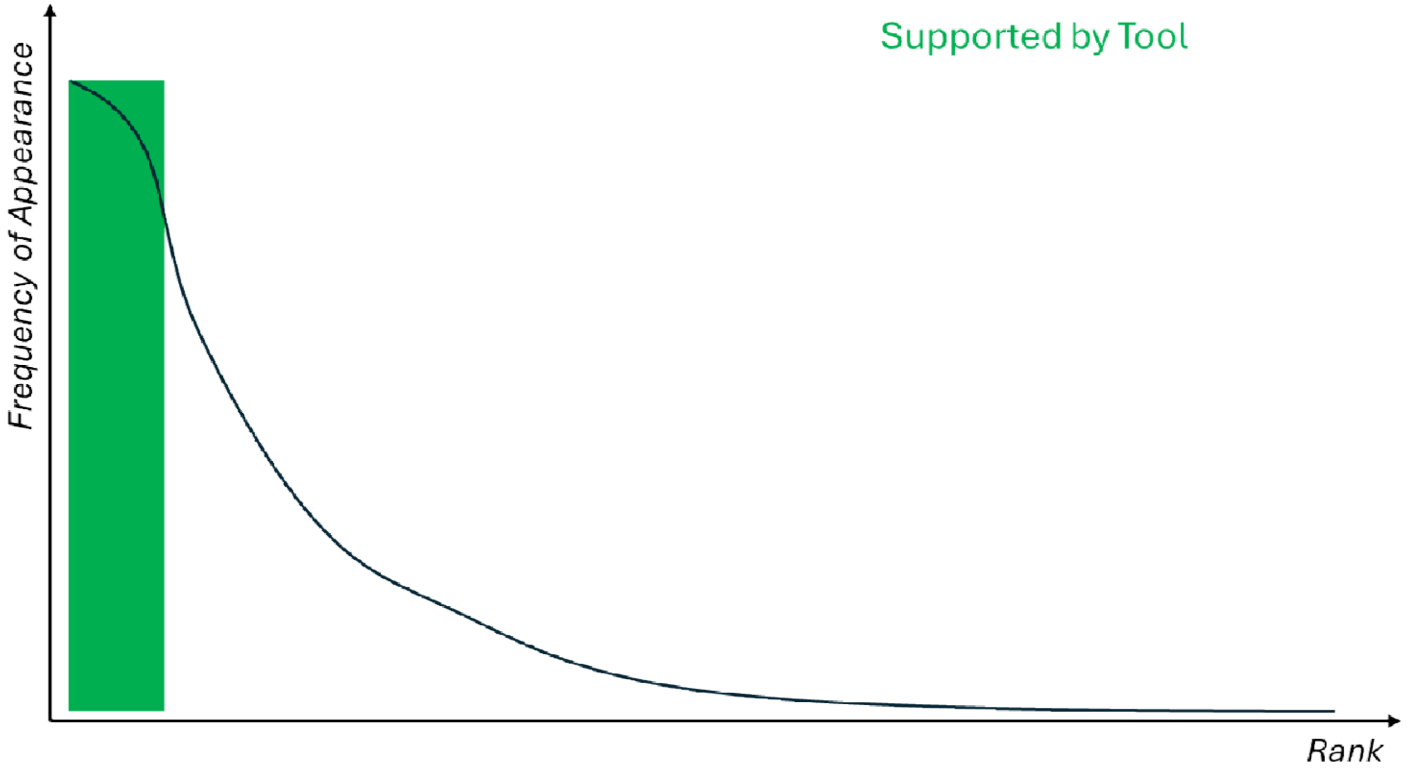
db_im_xx. The db_im_xx database keeps data about each users with whom userID has interacted with. It comprises solely of two tables, with only one, SIMPLE_USER, providing meaningful data. The SIMPLE_USER has as primary key the TikTok ID. For each listed user, the most relevant data are the NICK_NAME, the AVATAR_THUMB which holds JSON-formated content of the user's avatar including an URL to the thumbnail, UNIQUE_ID which represents the name handle and FOLLOW_STATUS. This last field stores the relationship between userID and the listed user: =0 does not follow but can be followed by userID, =1 follows and =2 follows and is followed, that is a *friend* in TikTok's parlance.

Domingues, P., Nogueira, R., Francisco, J.C., Frade, M., 2020. Post-mortem digital forensic artifacts of TikTok Android App, in: Proceedings of the 15th International Conference on Availability, Reliability and Security. Presented at the ARES 2020: The 15th International Conference on Availability, Reliability and Security, ACM, Virtual Event Ireland, pp. 1–8. https://doi.org/10.1145/3407023.3409203

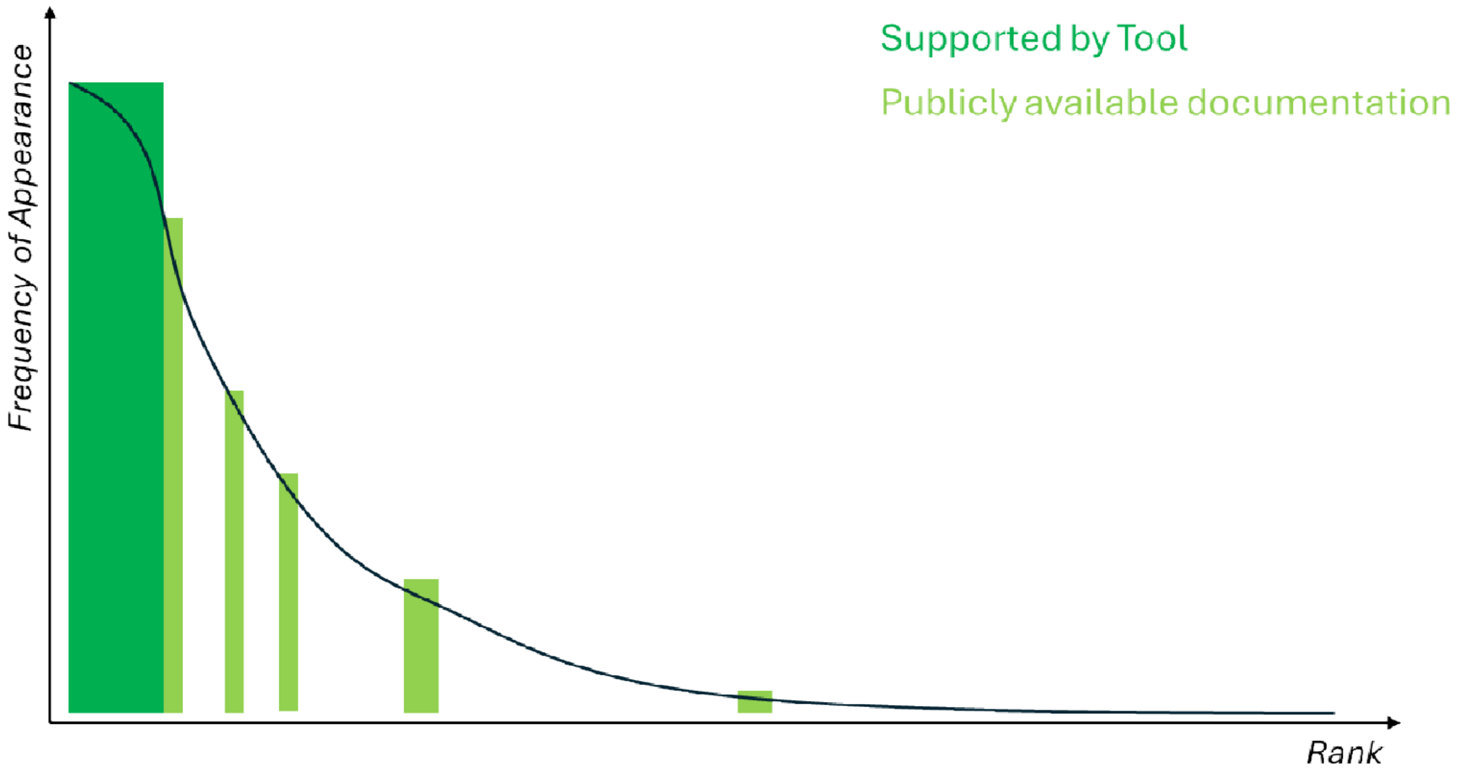
What Apps are documented?



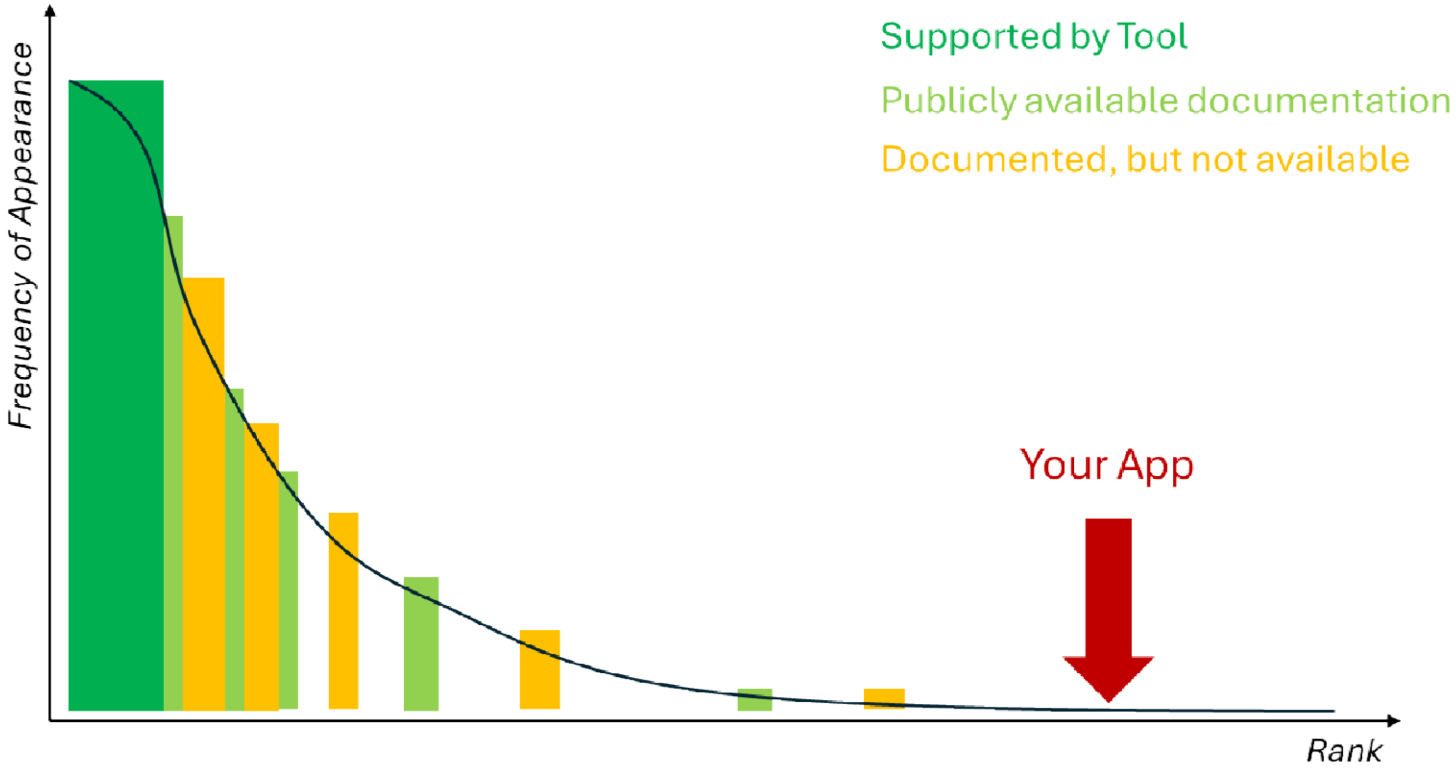
What Apps are documented?



What Apps are documented?



What Apps are documented?



Firefox: 10 releases in ~3 months (well documented *)

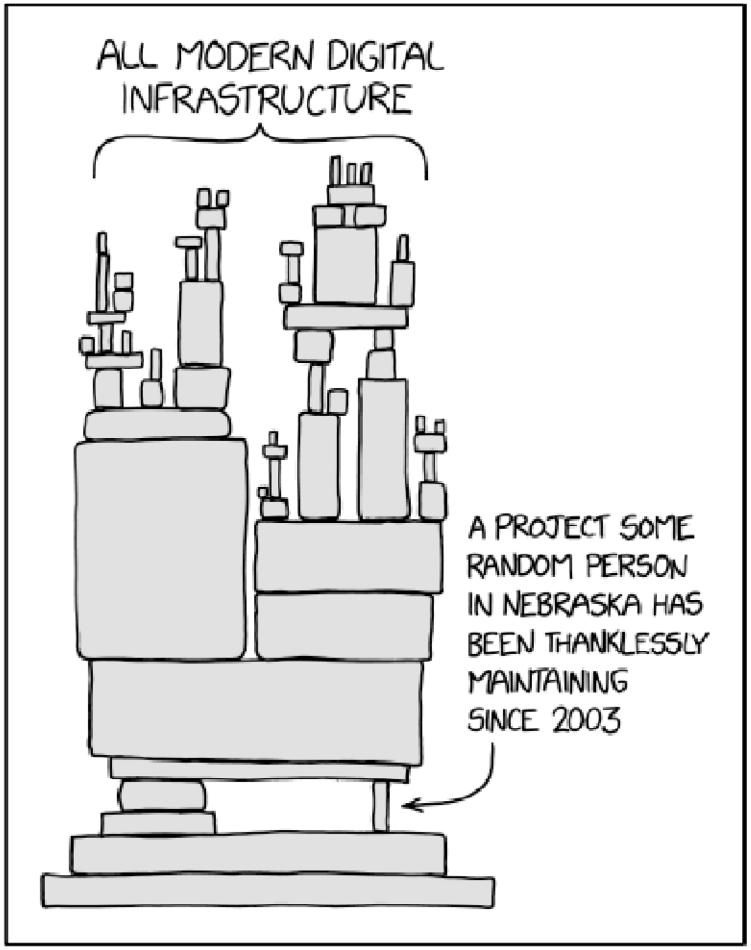
Version	Date	Release type
33.0	October 14, 2014	Official 33.0 release
33.0.1	October 24, 2014	Off-cycle stability update
33.0.2	October 28, 2014	Off-cycle stability update
33.0.3	November 6, 2014	Off-cycle stability update
33.1	November 10, 2014	**FF 10 year anniversary**
33.1.1	November 14, 2014	Off-cycle stability update
34.0	December 1, 2014	Official 34.0 release
34.0.5	December 1, 2014	Official 34.0.5 release
35.0	January 13, 2015	Official 35.0 release
35.0.1	January 26, 2015 (desktop) February 5, 2015 (Android)	Off-cycle stability update

* Source: https://software.**fandom**.com/wiki/Mozilla_Firefox_Version_History

Back to Dependencies

Every box in this picture has its own table of versions and releases.

But not all boxes have a fan club archiving version metadata.


ALL MODERN DIGITAL INFRASTRUCTURE
A PROJECT SOME RANDOM PERSON IN NEBRASKA HAS BEEN THANKLESSLY MAINTAINING SINCE 2003

https://xkcd.com/2347

How do we address this?

Please tell us if you know

Things we don’t think will work

- Just work harder
- Immediately gather reference data for everything
(or just everything you *will* need)
- Panic and give up
- AI

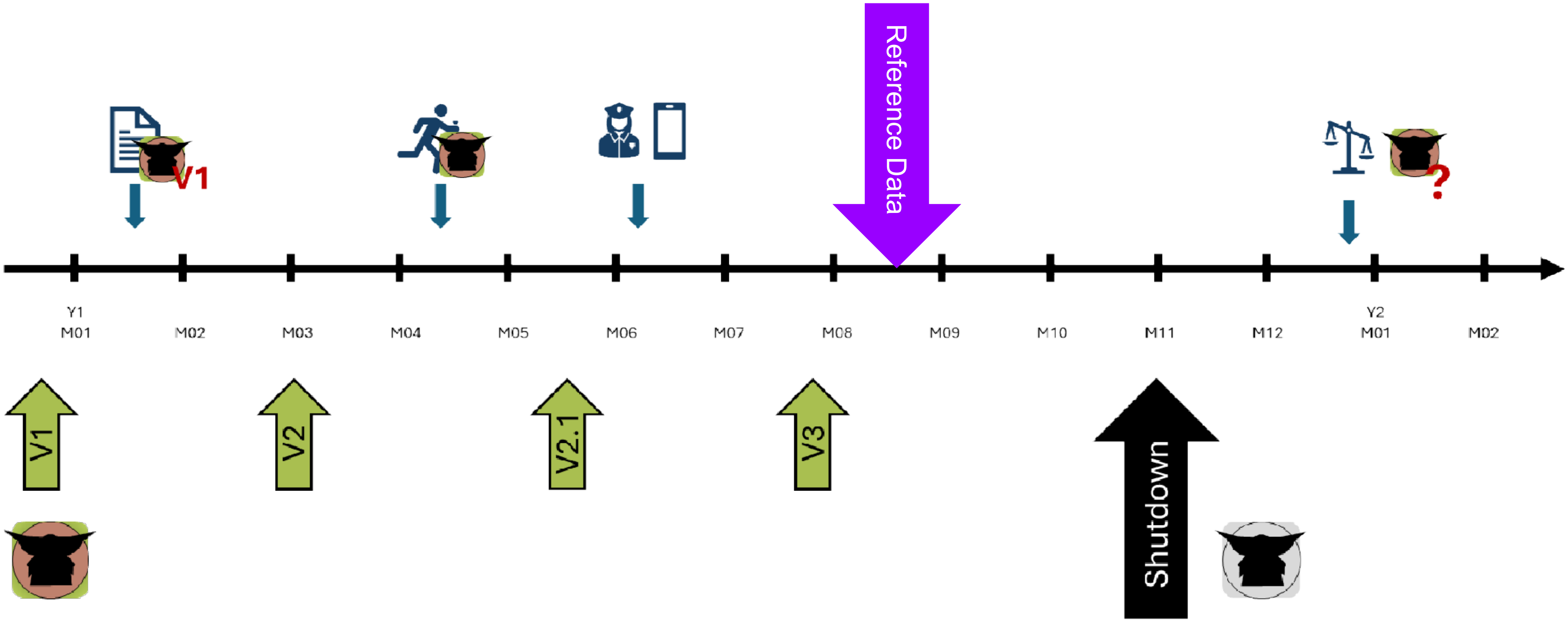
Updated Definition of Preservation

Preservation: protect traces from alteration (e.g., isolating them from surrounding environment), collect traces in a manner that changes as little as possible, and evidence management activities such as storing evidential items.

This includes gathering relevant supplemental information about the traces such as metadata, reference data, and context.

Adapted from: *Pollitt et al, 2018. A Framework for Harmonizing Forensic Science Practices and Digital/Multimedia Evidence, OSAC Technical Series. OSAC.*

Investigating a mobile application



Should we gather reference data for a given trace?

- How well is the service known and documented?
- How volatile is the system?
- How likely is the trace going to be used as evidence?
- Would exclusion of this piece of evidence massively impact the case?
- How much uncertainty is acceptable?
- …

Ideas to address this

- Public Database
  - Think: Artefact Genome Project + CFReDS + VirusTotal + Archive.org for forensic artifacts
  - Requires user contributions and maintenance
- Tool for automatic gathering of reference information
  - Document APIs, versions, watch for changes, try to link input and output
  - Watch and record IP addr/domain name changes
  - Investigators submit what to watch
- Taxonomy to assess risk and urgency
  - Qualitative rather than quantitative
  - Risk assessment tool: assess uncertainty in a case, provide suggestions to manage it
- None solve all problems, but some might help
- Must assess the feasibility of each of these

Summary

- The rate of change of software is increasing
- Remote servers and services change (and disappear) over time
- This makes testing unknown or poorly documented software challenging
- Testing is needed to understand the meaning of the client side trace data
- To mitigate this, we need research in new (feasible) approaches
- This isn’t the first time we’ve had to adapt to changes, it won’t be the last

Thank you for your attention

Hannes Spichiger (HSLU), hannes.spichiger@hslu.ch
Frank Adelstein (Hexordia), frank@notfrank.com

HSLU Hochschule Luzern

HEXORDIA

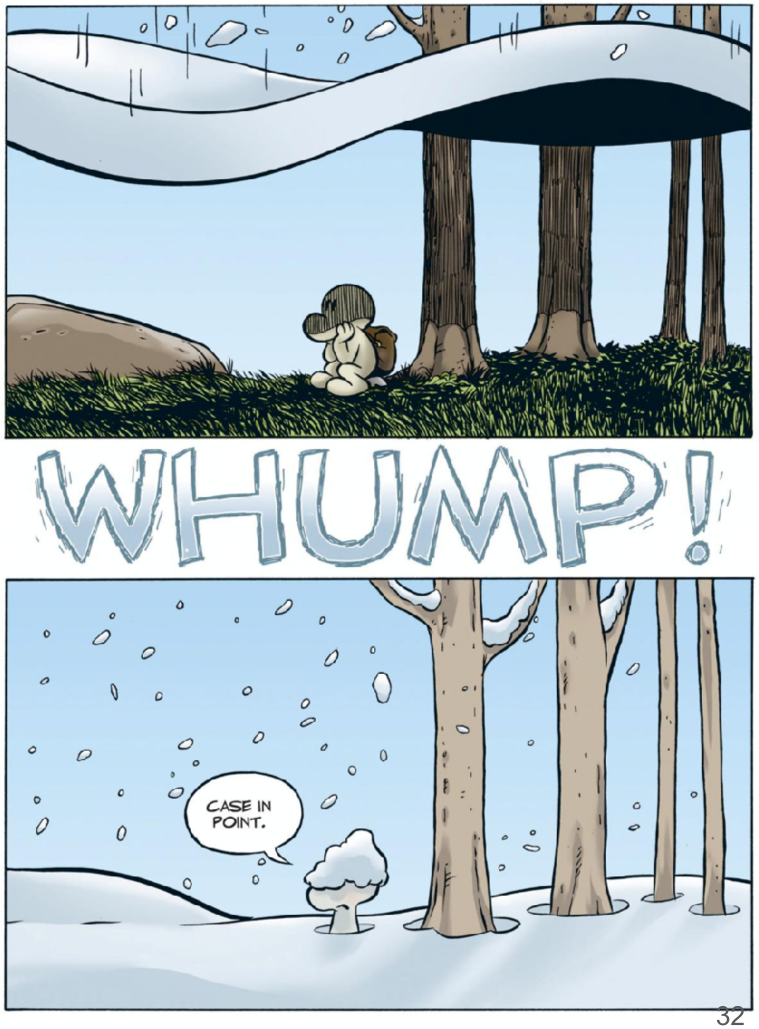
The End of the Golden Era of Digital Forensic Science…Again

2010: Garfinkel: *Digital forensics research: The next 10 years.*

2016: Quick et Choo: *Big forensic data reduction: digital forensic images and electronic evidence.*

2022: Pawlaszczyk: *Mobile Forensics – The End of a Golden Age?*

2025: Spichiger et Adelstein: *Preserving Meaning of Evidence from Evolving Systems*



Jeff Smith, 1991. *Bone*, Issue #1

The End of the Golden Era of Digital Forensic Science…Again

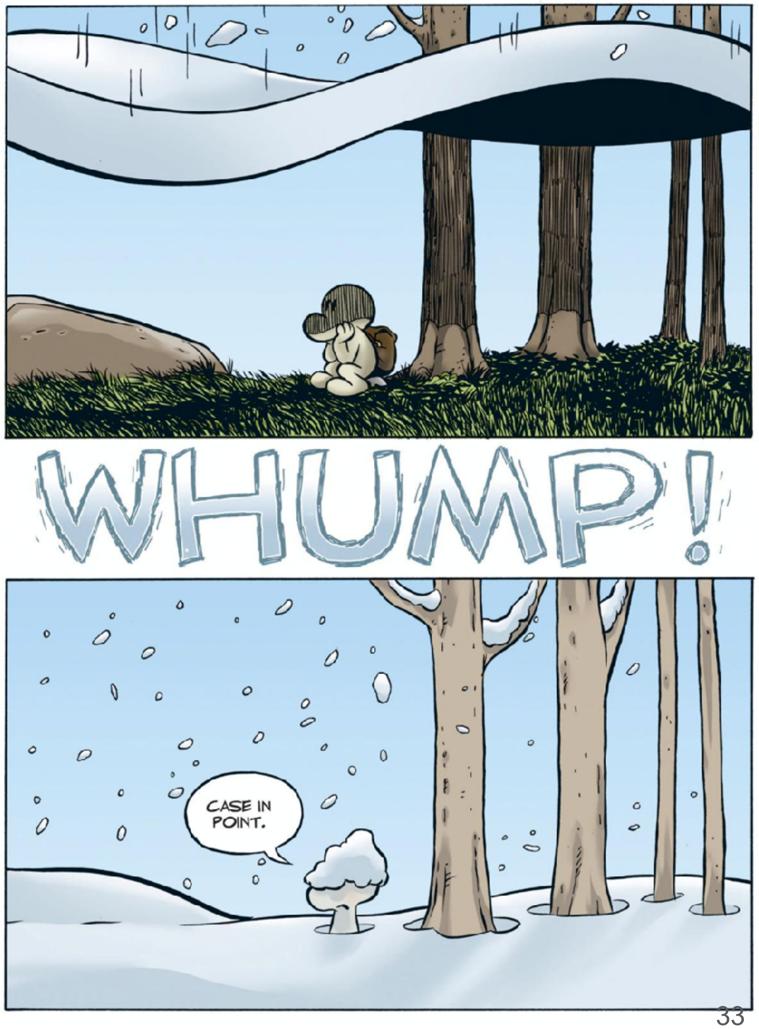
2005: There is too much data

2010: TRIM will be the end of data recovery

2015: Everything will be encrypted

2020: Mobile devices are too well protected.

2025: Rapidly evolving distributed system



Jeff Smith, 1991. *Bone*, Issue #1